DOCUMENT RESUME

ED 090 907 IR 000 157

AUTHOR	Mathews, Walter M.
TITLE	Humanizing Assessment Reports with a Computer.
PUB DATE	Apr 73
NOTE	13p.; Paper presented at the Annual Convention of the Association for Educational Data Systems (New Orleans, Louisiana, April 1973)
EDRS PRICE	MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS	*Computer Oriented Programs; Computers; *Evaluation; Humanism; *Humanization; *Narration; Speeches; Testing; *Test Results
IDENTIFIERS	AEDS; *Association for Educational Data Systems

ABSTRACT

Five computerized narrative assessment reports are discussed. These are: (1) the Teaching Information Processing System Student Report, used for a college economics course; (2) the Preliminary Scholastic Aptitude Test (PSAT) Score Report; (3) the Programmed Composition of Psychological Test Reports employed at the Mayo Clinic for reporting results of the Minnesota Multi-phasic Personality Inventory and the California Psychological Inventory; (4) a report devised by the Madison, Wisconsin public schools to explain individual performance on the Iowa Tests of Basic Skills; and (5) a report developed under the auspices of the Iowa Testing Programs to serve as a model narrative system for communicating results of the Iowa Tests of Basic Skills. The conclusion is reached that such narrative reports have several advantages, including flexibility, efficiency, and clarity of communication. In addition, since they provide personalized and self-explanatory information to students, parents, and other laymen, they serve an important public relations function. (PB)

HUMANIZING ASSESSMENT REPORTS WITH A COMPUTER

Walter M. Mathews

The University of Mississippi

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 090907

INTRODUCTION

When a person takes a test (or in some way is formally assessed), a report is usually produced. The report may be simply the number of correct items on the test, or it may be a detailed description of the testee's responses to a psychological test battery with comments and interpretation from a psychologist. The purpose of assessment reports is communication of information for decision-making, and the communication must be understandable to the receiver of the report.

Assuming no difficulties with the validity and reliability of the test, and a general understanding of the purpose and method of the test, the major barrier to understanding a performance report is understanding the language that is used. The language is usually numerical values of quantitative concepts such as raw scores, grade-equivalent scores, stanines, percentiles and standard errors of measurement. Certainly the target audience of the reports (students or parents or instructors or counselors, etc.) determines the kind of concern that language problems elicit; but to err by assuming an undistorted transmission of information is easy and a common occurrence.

The result of an understandable report should be the transfer of information on performance that can be used in the decision-making process. In the case of standardized achievement tests in education, Goslin found (3, p. 32) that elementary teachers receive score reports on pupil performance about 80 percent of the time and have free access to these scores in virtually every case, yet this information is shared routinely with pupils and their parents less than eight percent of the time. Current test reporting practices, then, seem to indicate that at least some of the information that could be useful in making educational decisions about a child is not routinely shared with the child and his parents. Proceeding on the assumption that testing information should be shared with parents and children, the obvious prohibiting factor is that "they would not understand the report." That is probably true. After all, some of our teachers and administrators have difficulty understanding current test reports. An alternative to depriving parents and students of testing information is to produce a testing report that is understandable to them. The suggestion and theme of this paper is the use of reports that are in a narrative format -- words that blend to form sentences which join into paragraphs.

DR. WALTER M. MATHEWS is assistant professor of educational administration at the University of Mississippi, teaching courses in computer applications, systems analysis, research design and statistics. He received his Ph. D. from the University of Wisconsin in 1971.

TM 000 157

Let's look at a few early attempts to utilize the computer in generating narrative reports.

Teaching Information Processing System (TIPS)

Kelley (5), a professor of economics at the University of Wisconsin, developed a program, which he called TIPS, to assist him in teaching. TIPS involved periodic collection of information from students regarding either their understanding of course materials or their reaction to various aspects of course presentations. TIPS provided a means of efficiently utilizing this information for instructional purposes. The information, which is collected on specialized forms suitable for machine processing, is composed of student responses to a series of multiple-choice questions. Surveys of six to twelve questions took about five to ten minutes to administer. Within a few hours this information was processed and summarized in three separate reports: one for each student, one for each section leader, and a third for the professor.

The student report contained a summary of his performance: his response to each question, the correct answers, and the total number of his correct answers. On the basis of this information, assignments for the forthcoming period were also indicated. The assignments (some required -- some optional) varied considerably in nature, level, and intensity. A student scoring well might have received optional assignments and/or required work at a higher level. The student performing poorly might have received not only the required work, but also a set of materials designed to bring him toward the mean class performance.

Additional information on the student report was generated on the basis of past as well as current performance. If the student had performed poorly over several surveys, he was instructed on the student report to establish an appointment with the instructor or teaching assistant. If the student performed consistently and exceptionally well, he might have been notified that a short paper could be substituted -- at his option -- for the midterm examination. A sample of a TIPS student report is presented as Figure 1.

The teaching assistant report contained information to help him appraise the performance of his individual sections, including statistics on percentage correct by question or by concept, actual responses on the survey, lists of students required to establish appointments or tutorials, and so forth.

The professor's report was similar to that received by the teaching assistant, although the information available applied to all students enrolled in the course rather than only to the students in particular sections. With this information the professor might elect to alter lectures, section coverage, problem sets, or other teaching instruments for the forthcoming period. In summary, TIPS is a system for gathering and reporting objective and timely information useful for more effective teaching.

Figure 1

SAMPLE COPY OF KELLEY'S TEACHING INFORMATION PROCESSING SYSTEM (TIPS) STUDENT REPORT

TIPS
STUDENT PERFORMANCE SURVEY RESULTS
PRINCIPLES OF ECONOMICS (103)
PROFESSOR ALLEN C. KELLEY

TURNER, LAWRENCE
715648
SURVEY TAKEN 10/25/67
SECTION NUMBER AND TIME 2, 9:55 F
SECTION LEADER MISS GREEN

OUT OF A TOTAL OF 10 QUESTIONS, YOU CORRECTLY ANSWERED 3. THE QUESTION NUMBER, YOUR RESPONSE AND THE CORRECT ANSWER ARE PROVIDED IN THE TABLE BELOW. YOU ARE URGED TO MAKE SURE YOU UNDERSTAND THE NATURE OF ANY INCORRECT RESPONSES YOU MAY HAVE MADE.

SUMMARY OF SURVEY RESULTS

QUES. NUM.	YOUR ANSW.	CORR. ANSW.	QUES. NUM.	YOUR ANSW.	CORR. ANSW.
1	G	G	6	B	A
2	C	B	7	D	E
3	D	B	8	A	A
4	A	C	9	D	C
5	F	F	10	C	B

YOUR ASSIGNMENT FOR THE WEEK, TO BE HANDED IN DURING THE DISCUSSION SECTION ON 11/03/67, IS THE FOLLOWING --

PROBLEMS 1, 3 AND 4 ON HANDOUT 2C

ADDITIONALLY, YOU ARE REQUIRED TO WORK THROUGH CHAPTER 2 OF - MICRO-ECONOMICS, A PROGRAMMED BOOK, BY LUMSDEN, ATTIYEH, AND BACH. IT WOULD BE USEFUL TO CONSULT THE PROGRAMMED UNIT BEFORE YOU READ HANDOUT 2C.

TO TEST YOUR UNDERSTANDING OF THE MATERIALS, YOU MAY, AT YOUR OPTION, ELECT TO COMPLETE THE FOLLOWING PROBLEMS --

1. WORKBOOK, PP. 37-38.
2. HANDOUT 2A, PROBLEMS 3 AND 4.

NOTICE -- AN ALL UNIVERSITY LECTURE BY PROFESSOR MILTON FRIEDMAN, PAST PRESIDENT OF THE AMERICAN ECONOMIC ASSOCIATION, WILL BE HELD IN 6210 SOCIAL SCIENCE, 7:30 P.M., 11/01/67. THE TOPIC -- MONETARY AND FISCAL POLICY RECONSIDERED.

Preliminary Scholastic Aptitude Test (PSAT) Score Reports

The effective reporting of test results to admissions officers, guidance counselors, and the individuals who take educational tests is a rather more complex matter than might appear at first glance. It is apparent that reporting only numerical scores is hardly adequate. The statistical, psychological, and educational contexts which allow the user to infer relevant meanings must be provided as well. The problem is particularly acute for the programs in which the primary reporting target is the test taker himself.

Since the precise interpretation of mental test scores does require a knowledge of statistics, measurement, psychology, and education, it is common practice to have test scores reported to individual test takers by guidance counselors who are expected to have the necessary sophistication. Unfortunately, not all guidance counselors do have sufficient psychometric sophistication, nor do they have the time to prepare detailed analyses and to give individual interpretation of mental tests results. Of course, most educational testing services prepare a wide range of interpretive materials for their testing programs to aid both the counselor and the student in interpreting test scores. But even with these aids, some statistical sophistication is still needed for adequate understanding.

Helm and Harasymiw (4) designed variable format computer-generated letters to be sent to examinees of the Preliminary Scholastic Aptitude Test as a report of their performance on the test. They found it necessary to prepare eighty distinct sentences where a sentence might have variable phrases imbedded. These eighty sentences were used to compose seventy-five distinct paragraphs which, in turn, were combined to produce the 100 letters needed to interpret all combinations of verbal and mathematical scores. They wrote a computer program to generate these letters and a sample of their report appears as Figure 2.

Programmed Composition of Psychological Test Reports

A computer-generated verbal diagnostic report on a standardized psychological test has been used routinely at the Mayo Clinic in Rochester, Minnesota since 1962 (7) and is available commercially from Behaviordyne, Inc. The Mayo program utilizes the Minnesota Multi-phasic Personality Inventory (MMPI) which is an objective pencil-paper psychological test. The machine-produced report is a group of disconnected statements, or decisions, about the subject as measured by the scales of the MMPI. Finney improved the Mayo program by offering an alternative of adding scales from the California Psychological Inventory (CPI), and by improving the coherence of the report (2). A large number of scales are scored, 101 with the MMPI alone or 124 with the MMPI and the CPI.

The report is built by selecting statements and then combining them into paragraphs. For each of the 101 scales, one statement is chosen from among eight possible statements, depending on the individual's score on the scale. By this method, 101 statements are chosen from a repertory of 808. Finney's program has the computer compose a full report on each individual's personality -- the kind of report that a psychologist might write after seeing a person several times and administering a full battery of tests. The first one-third of a sample report is given as Figure 3.

Figure 2

BEST COPY AVAILABLE

SAMPLE COPY OF THE PRELIMINARY SCHOLASTIC
APTITUDE TEST (PSAT) SCORE REPORT

EDUCATIONAL TESTING SERVICE

PRINCETON, N.J. 08540

JANUARY 21, 1966

DEAR MR. LETR 99,

WE WANT TO REPORT TO YOU THE SCORES YOU EARNED ON THE PRELIMINARY SCHOLASTIC APTITUDE TEST YOU TOOK ON OCTOBER 9, 1965. YOUR APTITUDE FOR COLLEGE WORK IS OUTSTANDING. IF YOUR HIGH SCHOOL MARKS ARE CONSISTENT WITH THE HIGH SCORES YOU HAVE EARNED ON THE TEST YOU WILL HAVE LITTLE DIFFICULTY IN BEING ACCEPTED AT A COLLEGE OF YOUR CHOICE.

YOU EARNED A SCORE OF 68 ON THE VERBAL SECTION OF THE TEST. THE VERBAL SECTION OF THE TEST MEASURES YOUR ABILITY TO READ WITH UNDERSTANDING AND TO USE WORDS EFFECTIVELY. A SCORE AS HIGH OR HIGHER THAN THE ONE YOU HAVE EARNED IS EARNED BY LESS THAN 11 PER CENT OF JUNIORS OF YOUR SEX WHO LATER ENTER COLLEGE. VERBAL APTITUDE IS PARTICULARLY IMPORTANT FOR SUCCESSFUL COLLEGE WORK IN THE HUMANITIES AND FINE ARTS.

YOU EARNED A SCORE OF 68 ON THE MATHEMATICAL SECTION OF THE TEST. THE MATHEMATICAL SECTION OF THE TEST MEASURES YOUR ABILITY TO REASON AND WORK EFFECTIVELY WITH NUMBERS. A SCORE AS HIGH OR HIGHER THAN THE ONE YOU HAVE EARNED IS EARNED BY LESS THAN 11 PER CENT OF JUNIORS OF YOUR SEX WHO LATER ENTER COLLEGE. MATHEMATICAL APTITUDE IS PARTICULARLY IMPORTANT FOR SUCCESSFUL COLLEGE WORK IN THE SCIENCES AND ENGINEERING.

YOU SHOULD NOT THINK OF YOUR TEST AS EXACT POINTS BUT AS A RANGE OF SCORES EXTENDING ABOUT THREE POINTS ABOVE AND THREE POINTS BELOW THE SCORE WE HAVE REPORTED TO YOU. THEY GIVE A GOOD INDICATION OF HOW YOU MAY EXPECT TO SCORE ON THE SCHOLASTIC APTITUDE TEST. THE CHANCES ARE FOUR OUT OF FIVE THAT YOU WILL SCORE BETWEEN 630 AND 730 ON THE VERBAL SECTION OF THE SCHOLASTIC APTITUDE TEST WHEN YOU TAKE IT NEXT YEAR. THE CHANCES ARE FOUR OUT OF FIVE THAT YOU WILL SCORE BETWEEN 630 AND 730 ON THE MATHEMATICAL SECTION OF THE SCHOLASTIC APTITUDE TEST WHEN YOU TAKE IT NEXT YEAR. INSOFAR AS VERBAL AND MATHEMATICAL APTITUDES ARE CONCERNED YOU CAN HAVE CONFIDENCE IN YOUR ABILITY TO DO SUCCESSFUL COLLEGE WORK.

THE SCORES YOU HAVE EARNED SHOULD ENCOURAGE YOU TO APPLY FOR ADMISSION TO AN OUTSTANDING COLLEGE. YOUR FUTURE EDUCATIONAL PLANS SHOULD CONSIDER ADVANCED GRADUATE WORK. AFTER YOU HAVE DISCUSSED YOUR SCORES WITH YOUR PARENTS AND YOUR COUNSELOR OR PRINCIPAL SHOULD YOU HAVE ADDITIONAL QUESTIONS ABOUT THEM YOU MAY WRITE TO EDUCATIONAL TESTING SERVICE, PRINCETON, N. J.

SINCERELY YOURS,

E. T. S.

Figure 3

SAMPLE COPY OF FINNEY'S REPORT ON THE MINNESOTA MULTI-PHASIC PERSONALITY INVENTORY

This is a report of MMPI and CPI testing of a female age 16, case number B-000005. This test, like any test, is subject to error. Testing only supplements other diagnostic examinations.

First let us examine the evidence of validity and the attitude with which she took the test.

RA-0, Cm-4
Gbs-10
L-3
Gi-4
Wb-3

On the CPI she gives mostly the common and conventional answers. That may be a sign of at least average common sense and judgment, and of being sufficiently steady, reliable, and realistic. She does not give a consistently favorable nor a consistently unfavorable picture of herself. She is afraid to admit even small flaws in herself, in terms of standards which are naive, rigid, perfectionistic, moralistic, unrealistic, and overly conventional. That shows a lack of insight. It also shows that she does not distinguish clearly between fundamental obligations, which people can and will expect her to meet, and the lesser or shallower matters in which a falling short in performance is tolerable. She wants to make a good impression in taking the test, and she gives the impression of having at least an average degree of warmth. She is moderately ambitious, alert, and productive, and likes working. She has no serious doubts about herself.

A-5, 7-4°
R-4, Rbw-5°
B-4°, Cfi-5
3f-27

She does not tell of anxiety or stress and is not looking for help. She is a reasonably compliant person. She has a normal amount of flexibility.

In terms of these factors, she seems to be a normal, average, flexible person.

2pc-31

But the two-point code tells us as follows. She has hysterical conversion reactions of some specific location or other. She is naive, exhibitionistic, self-centered, and demanding, and tends to manipulate and exploit people. Because of repression she lacks insight and is not motivated for psychotherapy.

Now, what is the evidence for psychosis or mental illness?

Rec
Scpt-3
Sx-3
43-7

None of the measures indicate that she is psychotic. Some measures are doubtful, as follows. The obsessive and schizophrenic indicators are about at equal level. On a schizophrenic correction scale, she seems to have psychotic trends. But most measures indicate that she is not psychotic. She does not use rituals or compulsive acts at all to ward off anxiety.

Next we consider narcissism, guilt, and basic trust.

N-7
Se-5
29-6
Gu-6
27-4
Em-7, 4-5°
Dom-4, Ex-5
6-3

Her self-esteem is low and she doesn't feel proud of herself. But by another measure her self-acceptance is within the average range, though she tends to blame herself a little more than the average. She shows signs of less than average guilt feeling. And she tends to deny guilt. She has a normal amount of concern with what people think of her. But she denies any feeling of self-consciousness or embarrassment. She has at least average dominance and initiative. She scores almost average on ego strength, and has fair tolerance for frustration. This is a good level of ego strength for a psychiatric patient. She has the assets of benefit from psychotherapy, but only if motivation and distress are also present. To a moderate degree she maintains an optimistic attitude by denying discouragement. She tells of very little worrying; less than the average person.

PtK-4°
44-4°, Ob-6
42-3, Ph-6

She shows signs of having some fears or phobias. But she does not admit fears or phobias.

Now we turn our attention to problems of dependency.

1-4
2-4
Rfd-5°, 3-4°
Ufl-5°, Mfy-4°

The signs are that she has only slightly more dependency need than the average, if at all so. Within the average or normal range she seems to put her dependency needs into action.

Now, what about being demanding or orally aggressive?

Finney and Auvenshire have subsequently developed several different kinds of reports written for different purposes. They are now extending their work to other objective psychological tests.

Baker suggested (1) that electronic computer technology be utilized to generate reports on other types of testing instruments in order to make the results more meaningful to the persons examined and to facilitate better use of the results of testing. He concluded:

> The mechanics of having the computer program prepare verbal descriptions depends upon several factors. First, the insight of the test constructor into the area of interest; second, the relation of levels of test and diagnostic scores to pupil performance; third, the cleverness of the computer programmer in generating connecting prose from somewhat disconnected verbal descriptions.

The Madison Project

The purpose of the Madison Project (6) was to design, produce, and evaluate computer-generated testing reports that would be useful to teachers and informative to parents, without requiring either parents or teachers to spend a lot of time preparing the reports or learning how to interpret them. In both cases, the reports were in a narrative format. A unique report was generated by a computer for each pupil, based on his scores on the fourth-grade level of the Iowa Tests of Basic Skills.

Attempts were made: to increase the communicative role of the reports by putting them in a narrative format; to personalize the reports by lacing each report with the pupil's name and the appropriate personal pronouns; to make the reports more understandable by eliminating or deemphasizing certain interpretative concepts (viz. grade-equivalent scores and percentile ranks, and a basic form for the explanatory paragraphs was developed, and then several parallel forms of each paragraph were written to provide variety of style. The first five paragraphs of each report described the pupil's performance on the text battery as a whole and on each of the sub-tests. Next, lists of specific areas were printed for which the pupil was diagnosed as either proficient or deficient -- areas, in which he was neither, were suppressed. These decisions were based on item analyses performed by the teachers in the Madison (Wisconsin) Public Schools and criterion levels set by them.

Figure 4 is a sample of a report to parents. The teacher report was very similar to this report, with the addition of imbedded percentile ranks. The parent reports were distributed at a parent-teacher conference and were the basis for a discussion of the child's performance.

One other report was generated -- a class-summary report for the teacher. This report was similar in format to the report on an individual student, but it discussed the class as a whole and concluded with specific suggestions that might be useful to the teacher in improving class performance in the tested areas in which their performance was weakest.

Figure 4

PARENT REPORT ON INDIVIDUAL PERFORMANCE ON THE IOWA TESTS OF BASIC SKILLS
ADMINISTERED SEPTEMBER 1970
MADISON PUBLIC SCHOOLS
MADISON, WISCONSIN

STUDENT: DANIEL MARSHALL
TEACHER: MRS. ANASTASI

ACCORDING TO THIS TEST, DANIEL'S OVERALL ACHIEVEMENT LEVEL IS AT THE 75 PERCENTILE WHEN COMPARED WITH STUDENTS IN GRADE 4 IN THE MADISON SCHOOL SYSTEM, AND AT THE 79 PERCENTILE WHEN COMPARED WITH A NATIONAL SAMPLE OF GRADE 4 STUDENTS.

THE TEST PROVIDES INFORMATION IN FIVE MAIN AREAS: DANIEL'S VOCABULARY SCORE IS NEAR AVERAGE. IN THE AREA OF LANGUAGE SKILLS HIS SCORE IS QUITE STRONG, AND HIS SCORE IN WORK-STUDY SKILLS IS VERY STRONG. HIS ARITHMETIC SCORE IS EXCEPTIONAL.

THE TEST LOOKED AT FOUR SPECIFIC LANGUAGE SKILLS, AND THIS STUDENT WAS RATED AS FOLLOWS: QUITE STRONG IN SPELLING AND QUITE STRONG IN CAPITALIZATION. RATHER WEAK IN PUNCTUATION AND NEAR AVERAGE IN WORD USAGE.

IN THE AREA OF WORK-STUDY SKILLS, THREE SPECIFIC SKILLS WERE TESTED. IN MAP READING, DANIEL SEEMS TO BE QUITE STRONG WHILE HE IS NEAR AVERAGE IN READING GRAPHS AND TABLES. IN HIS KNOWLEDGE AND USE OF REFERENCE MATERIALS HE IS QUITE STRONG.

IN THE AREA OF ARITHMETIC, HE APPEARS TO BE NEAR AVERAGE IN CONCEPTS AND VERY STRONG IN PROBLEM SOLVING WHILE HE TENDS TO BE EXCEPTIONAL IN THE MODERN MATHEMATICS SUPPLEMENTARY TEST.

DIAGNOSTIC KEYS INDICATE THAT DANIEL IS PROFICIENT IN:

SOLVING PROBLEMS INVOLVING MONEY
SUBTRACTION.
ADDITION.
UNDERSTANDING THE CONCEPT OF WHOLE NUMBER IN MODERN MATH.

DIAGNOSTICS ALSO INDICATE THAT DANIEL IS DEFICIENT IN:

SPELLING BECAUSE OF PROBLEMS IN VOWEL SUBSTITUTION.
USING THE COMMA.
PUNCTUATION BECAUSE OF OVERPUNCTUATION

STUDENT PERFORMANCE IS EVALUATED IN SEVERAL WAYS. BY THE TEACHER, BY THE PARENTS OF THE STUDENT AND, OF COURSE, BY THE STUDENT HIMSELF. TEST RESULTS DO NOT REPLACE ANY OF THESE. BUT CAN BE VALUABLE IN THAT THEY ARE ANOTHER SOURCE OF DATA. THIS TEST, LIKE ANY TEST, IS SUBJECT TO ERROR.

COMMENTS ABOUT DANIEL MARSHALL FROM MRS. ANASTASI

Narrative testing reports were generated for a random half of fifty-two fourth-grade classes in seventeen schools. Three major conclusions were reached as a result of the valuation of the narrative reports when compared to the traditional reports: (1) The teachers preferred receiving the narrative reports to the traditional reports; (2) the teachers who received the narrative reports felt better prepared to communicate information obtained from the tests to the parents of their pupils than did teachers who received the traditional reports; (3) receipt of the narrative reports had a greater impact on the teacher's perceptions of pupil performance that did receipt of the traditional reports.

The Madison Project did a respectable job at what it did, but it did not do enough. It took a major step in humanizing testing reports and in offering an alternative to the obstacle of terminology and numerical ratings, but it only reported the pupil's performance compared to other fourth graders -- no reference was made to the child's performance compared to his own previous performance. The Madison reports generated the same basic format for all children, making no allowances for different needs at different levels of performance. The greatest shortcoming of these reports, however, was the absence of a pupil report.

The Iowa Project

A project conducted during the summer of 1972 under the auspices of the Iowa Testing Programs developed a model for a narrative reporting system for the fifth-grade level Iowa Tests of Basic Skills. The model built upon the Madison Project, but the emphasis was on the basic conceptualization rather than on developing the technology for implementing the model.

The most important changes from the Madison Project to the Iowa Project were: (1) the addition of a pupil report; (2) the inclusion of growth statements; and (3) the use of three alternative levels of the parent and pupil reports -- levels that are based on overall pupil performance on the test battery. An example of the kind of pupil report we developed as Figure 5.

The growth statements are verbal in nature and are based on the difference between the score obtained by the pupil on the current year's testing and the score obtained at some previous administration, usually one year earlier. The three alternative levels of the parent and pupil report approach their audiences from different bases: a low performing pupil is given encouragement after showing growth, even though his performance level is low, while a high-performing pupil may be given specific suggestions on how to improve in an area in which his performance is high, if it is his weakest area.

Other changes also were designed: the parent report was expanded to include a general explanation of tests and test results; suggestions for increasing performance were included in the parent and pupil reports; comments to reinforce growth were added; the reading level of the reports was lowered; the teacher report was shrunk to a skeletal report.

BEST COPY AVAILABLE

Figure 5

PUPIL REPORT ON PERFORMANCE ON THE IOWA TESTS OF BASIC SKILLS
Administered January 1973
Iowa City Public Schools
Iowa City, Iowa
HORACE MANN SCHOOL

Pupil: JOHN BLACK GRADE 6 Teacher: MRS. RICHARDSON

A few weeks ago, you took the Iowa Tests of Basic Skills. You may remember that it had questions on vocabulary, spelling, capitalization, punctuation, and reading maps, graphs, and tables. There were also questions about short stories you were asked to read, and there were problems in arithmetic, too.

These tests give you a chance to find out how you are doing in the skills that are used in school. Your teacher gave the tests to everyone in the class to help her see in what areas the class needs most work. Another reason for giving the tests was to give you information on your skills--which skills are strong, and which ones are weak. It's good to find out which skills need attention, since then you can work to improve them.

This report is to let you know how you did in the tests. The first comment for each test tells you how strong or weak your score was in that area, compared to the scores of other students in sixth grade in Iowa. Your scores is also compared to the score you had on a similar test you took in the fifth grade. This shows how much you grown in that area since last year.

VOCABULARY — In the test on vocabulary, your score was near average. Your growth during the past year was below average. One good way to improve your vocabulary would be to take a little time to look in the dictionary for the meaning of words you are not sure of. It takes extra time, and sometimes you might have to look up the same word several times before you remember its meaning. But your vocabulary will get bigger and you will better understand what you read and hear. And you will be better able to express what you think.

READING COMPREHENSION — Your performance on this test showed strength. You showed normal growth in reading in the past year. In this test you did very well in seeing and understanding important facts and details in what you read. You must like to read, because you do it so well.

LANGUAGE SKILLS — In the four tests that measured the mechanics of writing, your overall performance was near average. Your growth from last year was below average, and these are some areas where a little extra work would pay off.

SPELLING — Your score on this test was very weak. You showed minimum growth since last year. Looking a little closer at some of the questions you missed on this test, it appears that you sometimes are not sure which vowels correctly spell a word. If you really want to improve your spelling, you can. Use a dictionary to check spellings. Your teacher will give you help too, if you want it.

CAPITALIZATION — Your performance in capitalization was near average. You showed normal growth. You did very well in capitalizing geographic terms. Using proper capitalization in letter writing is an area in which you could use some work.

BEST COPY AVAILABLE

PUNCTUATION — Your performance in this area of language skills was less than average. Your growth was below average. You tended to use too much punctuation, and you may want to watch that. The use of punctuation in writing letters was notices as an area in which you made several mistakes. Mrs. Richardson can help you in these areas if you would like.

USAGE — Your performance in the proper use of words was near average. Your growth since last year was normal. In this test you did very well in using verbs.

WORK-STUDY SKILLS — There were three tests that measure work-study skills; a map reading test, a test of reading graphs and tables, and a test of your knowledge and use of reference materials. In general, your performance was near average. You showed normal growth.

MAP READING--Your score on this test was near average. Your growth since last year was below average. You seemed to have trouble locating places on maps and globes. Some practice in using a map of Iowa City and the state of Iowa would be helpful. (The attendant at a gas station might give you a map if you asked for it.)

READING GRAPHS AND TABLES--You have shown strength in this skill. Your growth in this area is above average. You've improved in this area and should be proud of it.

REFERENCE MATERIALS--Your knowledge and use of reference materials is near average, according to the test. You've shown normal growth since last year. You did miss questions on the use of an index. You might want to practice using the index that you find at the back of many books.

MATHEMATICS SKILLS — Your overall performance in the two tests of arithmetic skills shows you to have strength in this area. Your growth in these skills since last year was above average. In general, you are doing very well in math. Keep it up!

MATHEMATICS CONCEPTS--Your score on this test showed strength. You showed above average growth in this area.

MATHEMATICS PROBLEM SOLVING--You have shown strong performance in this area. Your growth in this skill was above average. You did particularly well in solving problems involving money. You did well in solving addition problems.

In general, these tests show that you have done very well in improving your mathematics skills in this past year, but not so well in vocabulary and language skills. You will want to pay close attention to spelling and punctuation. You also showed good progress in reading graphs and tables.

There are other things you do at school that are also important. You are growing in many ways, not only in your skills, but also as a person. The Iowa Tests of Basic Skills do not measure these other areas of growth, so they have not been talked about here. It's important that you , your parents, and your teacher talk about these other areas--your everyday activities in all areas of school, your interests, and how you feel about school. That was, all of you can be more sure that you are getting the most out of your experience in school.

Some Potential Advantages

Depending upon the specific application, several advantages accrue from the use of narrative reports. A few general advantages follow:

*Clear Communication -- The report is in understandable English and does not ask the recipient to search the page for clues on where to begin "reading" the report.

*Personal -- The use of the testee's name and the appropriate personal pronouns can easily be included in the text to humanize the report and increase the attention paid to it. It does not "look like a computer report."

*Efficient -- A more complete report is available at the cost of fewer professional-person-hours. Narrative reports are not suggested as a substitute for the professional, but rather as an aide to the professional.

*Self-Explanatory -- The report can be reasonably self-contained with less need for individual guidance to assist in its initial interpretation.

*Flexible -- The underlying philosophy of the narrative can be varied as needed for different audiences and audience sub-sets.

*Public Relations -- An inherent public relations value can accrue to the organization from a professional report that is understandable and personalized.

Certainly there are opposing arguments that could be offered:

*Muddled Communication -- The report has to be read, thereby restricting its audience. Realizing that, the reading level of the report still makes communication less than efficient with another segment of people. Others who can read well, do not want to be bothered to read -- they just want the score, and without extraneous verbiage.

*Too Personal -- Some people may react negatively to the thought of a machine attempting to be "intimate" with them.

*Expensive -- Initial development of the system is expensive with no promise of a lowered cost per report.

This list could also be extended, and perhaps you will add to it during the discussion. We do not pretend to propose a panacea for reporting-problems. It does appear, however, that narrative reports have a place, and that their application should be further explored to determine its precise location.

REFERENCES

1. Baker, Frank B., "Automation of Test Scoring, Reporting, and Analysis." Chapter 8 in Educational Measurement, 2nd ed., R.L. Thorndike ed. (Washington, D. C.: American Council on Education, 1971).

2. Finney, Joseph C., "A Programmed Interpretation of the MMPI and the CPI," Archives of General Psychiatry, XV (1966), pp. 202-36.

3. Goslin, David A., Teachers and Testing. (New York: Russell Sage Foundation, 1967).

4. Helm, Carl E. and Harasymiw, Stefan J., "Computer-based Score Reports," Measurement and Evaluation in Guidance, I, 1 (Spring, 1968), pp. 27-35.

5. Kelley, Allen C., "An Experiment with TIPS: A Computer-aided Instructional System for Undergraduate Education," The American Economic Review: Papers and Proceedings, LXXX (1968), pp. 446-57.

6. Mathews, Walter M. "The Development of Computer-generated Verbal-format Testing Reports," Journal of Educational Data Processing, IX, 3 (1972), pp. 1-11.

7. Swenson, W. M., et al., "Symposium on Automation Techniques in Personality Assessment," Proceedings of Staff Meetings of the Mayo Clinic, XXXVII (1962), pp. 61-82.